INSOLOGI: Jurnal Sains dan Teknologi
https://journal.literasisains.id/index.php/INSOLOGI
ISSN 2828-4984 (Media Online) | ISSN 2828-4992 (Media Cetak)
Vol. 4 No. 4 (Agustus 2025) 706-720
DOI: 10.55123/insologi.v4i4.5590
Submitted: 16-06-2025 | Accepted: 29-07-2025 | Published: 10-08-2025

Deteksi Dini Stroke Menggunakan Machine Learning

Kevinda Sari[1], Muhammad Fadli[2*], Muhammad Fahmi Fudholi[3], Erliyan Redy Susanto[4]
[1,3,4]Magister Ilmu Komputer, Universitas Teknokrat Indonesia, Kota Bandar Lampung, Indonesia
[2*]Jurusan Ekonomi dan Bisnis, Politeknik Negeri Lampung, Kota Bandar Lampung, Indonesia
Email: [1]kevinda_sari@teknokrat.ac.id, [2]muhammadfadliofficial@polinela.ac.id,
[3]muhammad_fahmi_fudholi@teknokrat.ac.id, [4]erliyan.redy@teknokrat.ac.id

Abstract

Stroke is one of the leading causes of death and disability worldwide. Early detection of stroke risk is crucial to prevent more severe complications. This study aims to develop a stroke prediction model based on machine learning using an open dataset from Kaggle containing patients' medical and demographic information. Four machine learning algorithms were utilized and compared: AdaBoost, Gradient Boosting, LightGBM, and XGBoost. Data preprocessing steps included missing value imputation, categorical variable encoding, numerical feature normalization, and class balancing using the SMOTEENN method. Additionally, feature selection was performed using the Extra Trees algorithm to enhance model performance. The results showed that the XGBoost model delivered the best performance, achieving an accuracy of 97.16%, an F1-score of 97.49%, and an AUC of 99.75%. This model proved to be effective in detecting stroke cases and holds potential for integration into clinical decision support systems. The study concludes that a combination of modern boosting algorithms and optimal preprocessing techniques can yield a reliable stroke prediction system suitable for implementation in digital healthcare contexts.

Keywords: *Stroke, Machine Learning, Xgboost, Early Detection, SMOTEENN, Medical Classification.*

Abstrak

Stroke merupakan salah satu penyebab utama kematian dan kecacatan di dunia. Deteksi dini terhadap risiko stroke menjadi sangat penting untuk mencegah komplikasi yang lebih serius. Penelitian ini bertujuan untuk mengembangkan model prediksi stroke berbasis machine learning dengan memanfaatkan dataset terbuka dari Kaggle yang berisi informasi medis dan demografis pasien. Empat algoritma pembelajaran mesin digunakan dan dibandingkan, yaitu AdaBoost, Gradient Boosting, LightGBM, dan XGBoost. Proses praproses data meliputi imputasi nilai hilang, encoding variabel kategorikal, normalisasi fitur numerik, serta penyeimbangan data menggunakan metode SMOTEENN. Selain itu, dilakukan seleksi fitur menggunakan algoritma Extra Trees untuk meningkatkan performa model. Hasil penelitian menunjukkan bahwa model XGBoost memberikan performa terbaik dengan akurasi sebesar 97,16%, F1-score 97,49%, dan AUC 99,75%. Model ini terbukti efektif dalam mendeteksi kasus stroke dan berpotensi untuk diintegrasikan ke dalam sistem pendukung keputusan klinis. Penelitian ini menyimpulkan bahwa kombinasi algoritma boosting modern dan teknik praproses yang optimal dapat menghasilkan sistem prediksi stroke yang andal dan dapat diimplementasikan dalam konteks layanan kesehatan digital.

Kata Kunci: Stroke, Machine Learning, Xgboost, Deteksi Dini, SMOTEENN, Klasifikasi Medis.

1. PENDAHULUAN

Stroke merupakan salah satu penyakit yang paling mematikan dan melemahkan di dunia (Swathi Priyadarshini and Abdul Hameed 2024). Menurut data dari Organisasi Kesehatan Dunia (WHO), stroke menempati peringkat kedua sebagai penyebab utama kematian secara global dan menjadi penyebab utama kecacatan jangka panjang (Zuama, Rahmatullah, and Yuliani 2022). Menurut (Indah Werdiningsih et al. 2023), stroke terjadi ketika pasokan darah ke otak terganggu, baik karena penyumbatan (stroke iskemik)

maupun karena pecahnya pembuluh darah (stroke hemoragik). Akibat dari gangguan tersebut adalah kematian jaringan otak yang dapat menyebabkan berbagai komplikasi, termasuk kelumpuhan, gangguan bicara, kehilangan memori, hingga kematian. Dalam konteks ini (Gupta et al. 2024), deteksi dini menjadi sangat krusial. Semakin cepat stroke dikenali dan ditangani, semakin besar peluang untuk menghindari kerusakan otak permanen dan menurunkan angka kematian (Salsabila et al. 2024).

Di sisi lain, Menurut (Putra et al. 2022) revolusi digital dan kemajuan teknologi informasi telah memberikan dorongan besar terhadap penerapan kecerdasan buatan (artificial intelligence/AI) dalam bidang kesehatan. Pembelajaran mesin (machine learning/ML), sebagai cabang dari AI, telah menunjukkan potensi yang luar biasa dalam berbagai aplikasi medis, termasuk prediksi penyakit, klasifikasi kondisi kesehatan, pengenalan pola dalam data medis, dan rekomendasi pengobatan. Dalam beberapa tahun terakhir, penerapan machine learning untuk prediksi penyakit kardiovaskular, termasuk stroke, telah menjadi fokus utama banyak penelitian karena kemampuannya untuk menangani data medis dalam jumlah besar dan kompleks.

Prediksi stroke secara tradisional bergantung pada diagnosis klinis yang dilakukan oleh tenaga medis berdasarkan gejala-gejala yang muncul dan pemeriksaan laboratorium. Namun, metode ini cenderung bersifat reaktif daripada proaktif. Pasien biasanya sudah mengalami gejala yang cukup parah sebelum akhirnya didiagnosis mengalami stroke. Oleh karena itu, menurut (Herlistiono and Violina 2024) pendekatan prediktif berbasis data menjadi penting untuk mengantisipasi kejadian stroke sebelum gejala klinis muncul secara nyata. Di sinilah machine learning memainkan peran penting. Dengan memanfaatkan data historis dari pasien yang terdokumentasi secara digital, model prediktif berbasis machine learning mampu mengenali pola-pola halus yang tidak kasat mata bagi manusia tetapi signifikan secara statistik.

Salah satu tantangan utama dalam prediksi stroke adalah kompleksitas faktor risiko yang terlibat. Stroke dapat disebabkan oleh berbagai kombinasi faktor seperti usia, jenis kelamin, tekanan darah tinggi (hipertensi), kadar gula darah, obesitas, riwayat merokok, pola makan, aktivitas fisik, genetika, serta kondisi medis lainnya seperti penyakit jantung atau diabetes. Kompleksitas ini menyulitkan pembuatan model prediktif sederhana yang hanya mengandalkan rumus atau aturan eksplisit. Machine learning memungkinkan pembuatan model non-linear yang mampu menangkap interaksi kompleks antar variabel dan menggeneralisasi pola dari data untuk membuat prediksi akurat (Chung et al. 2023).

Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi model prediktif stroke berbasis machine learning dengan memanfaatkan dataset terbuka dari Kaggle yang mencakup berbagai atribut medis dan demografis pasien. Peneliti fokus pada pengembangan model dengan akurasi tinggi untuk mengidentifikasi individu dengan risiko tinggi mengalami stroke. Beberapa algoritma machine learning digunakan dan dibandingkan dalam penelitian ini, antara lain AdaBoost (Adaptive Boosting), Gradient Boosting Machines (GBM), LightGBM (Light Gradient Boosting Machine), XGBoost (eXtreme Gradient Boosting). Peneliti memilih algoritma-algoritma ini karena terbukti memiliki performa yang baik dalam klasifikasi biner pada berbagai domain kesehatan.

Selain pemilihan algoritma, tahap praproses data juga merupakan komponen penting dalam membangun model prediktif yang efektif. Data medis umumnya mengandung banyak tantangan seperti data hilang (missing values), ketidakseimbangan kelas (class imbalance), dan keberadaan variabel kategorikal. Dalam penelitian ini, peneliti menerapkan sejumlah teknik praproses data termasuk imputasi nilai hilang, encoding variabel kategorikal, normalisasi fitur numerik, dan teknik SMOTEENN untuk menangani ketidakseimbangan kelas antara pasien stroke dan non-stroke. SMOTEENN

memungkinkan model untuk belajar dari data minoritas secara lebih efektif, sehingga meningkatkan sensitivitas model terhadap kasus-kasus stroke yang jarang terjadi.

Dalam proses evaluasi model, metrik utama yang digunakan adalah akurasi, sensitivitas, dan spesifisitas. Peneliti juga menggunakan confusion matrix untuk mendapatkan pemahaman yang lebih dalam mengenai performa masing-masing model, termasuk kemampuan model dalam mengidentifikasi kasus positif (stroke) dan negatif (non-stroke).

Penelitian ini juga membandingkan hasil dengan berbagai studi terdahulu yang menggunakan pendekatan serupa. Meskipun sebagian besar penelitian sebelumnya menggunakan algoritma populer seperti Random Forest, Naïve Bayes, CART, DNN atau KNN, pendekatan peneliti menunjukkan bahwa kombinasi teknik praproses yang optimal, pemilihan fitur yang relevan, serta penggunaan algoritma boosting dapat menghasilkan model prediktif dengan kinerja yang lebih tinggi. Hal ini menunjukkan pentingnya pendekatan holistik yang mencakup seluruh siklus pemodelan mulai dari pembersihan data hingga evaluasi performa model.

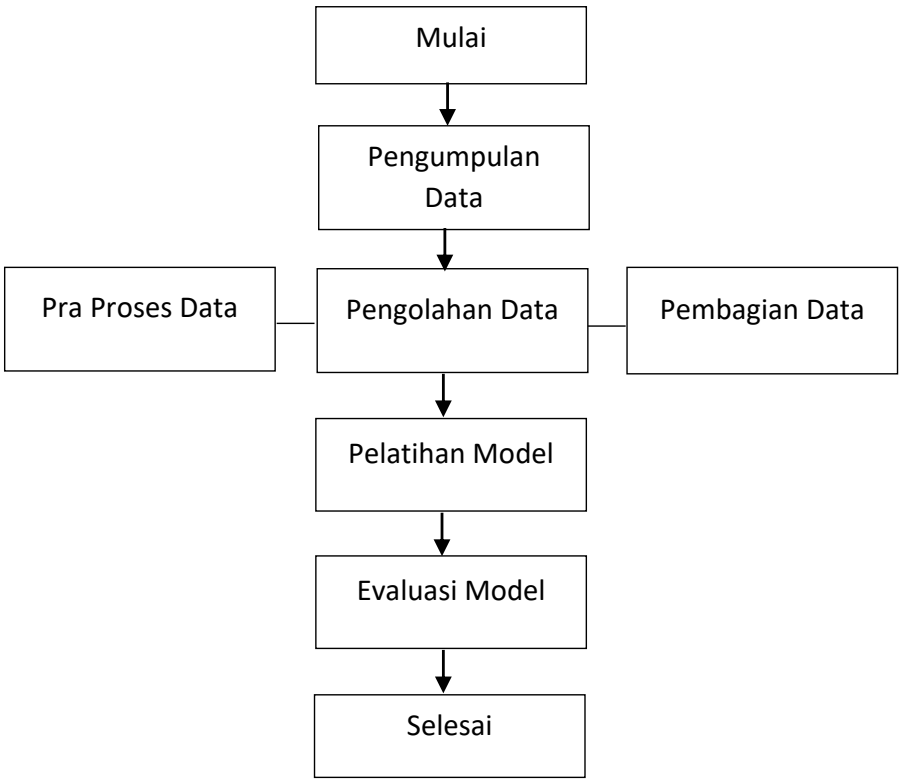
Kontribusi utama dari penelitian ini adalah penyajian pendekatan komprehensif untuk membangun model prediktif stroke dengan machine learning yang dapat digunakan sebagai alat bantu dalam sistem pendukung keputusan klinis. Dengan integrasi ke dalam sistem informasi rumah sakit atau aplikasi mobile berbasis kesehatan, model ini berpotensi untuk digunakan oleh dokter maupun pasien dalam melakukan penilaian risiko stroke secara cepat dan efisien. Penggunaan machine learning tidak dimaksudkan untuk menggantikan tenaga medis, melainkan sebagai alat bantu yang memperkaya proses pengambilan keputusan dengan informasi berbasis data.

Namun demikian, terdapat beberapa keterbatasan dalam penelitian ini yang perlu dicatat. Pertama, dataset yang digunakan bersumber dari sumber terbuka yang mungkin tidak sepenuhnya mencerminkan keragaman populasi atau kondisi klinis di berbagai wilayah geografis. Oleh karena itu, validasi lebih lanjut pada dataset lokal atau real-world clinical data diperlukan sebelum implementasi klinis yang luas. Kedua, data yang digunakan bersifat tabular dan statik, sehingga belum mencakup data waktu (time-series) yang mungkin memberikan informasi tambahan mengenai dinamika kondisi pasien. Ketiga, meskipun model yang dikembangkan menunjukkan performa yang baik, interpretabilitas model masih menjadi tantangan, terutama pada model kompleks.

Sebagai kesimpulan awal dari bagian pendahuluan ini, machine learning menawarkan pendekatan baru yang menjanjikan dalam bidang prediksi stroke. Dengan kemampuan untuk memproses dan menganalisis data dalam skala besar serta mengenali pola kompleks yang sulit ditangkap oleh manusia, algoritma machine learning dapat menjadi komponen penting dalam upaya pencegahan stroke secara proaktif. Penelitian ini memberikan kontribusi nyata dalam memperkuat literatur terkait penggunaan AI dalam prediksi penyakit, khususnya dalam konteks stroke, dan diharapkan dapat mendorong adopsi teknologi serupa dalam layanan kesehatan di masa mendatang.

2. METODOLOGI PENELITIAN

Untuk Penelitian ini bertujuan mengembangkan model prediksi dini stroke menggunakan algoritma pembelajaran mesin. Metodologi penelitian dilakukan melalui beberapa tahapan sistematis, mulai dari pengumpulan data, praproses data, seleksi fitur, pelatihan model, hingga evaluasi performa. Tahapan-tahapan tersebut dijelaskan pada gambar 1.



Gambar 1. Diagram Alir

2.1. Pengumpulan Data

Dataset yang digunakan bersumber dari Kaggle *https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset*. Dataset ini terdiri dari 5110 data pasien dengan 12 fitur, termasuk atribut seperti usia, jenis kelamin, hipertensi, penyakit jantung, status pernikahan, jenis pekerjaan, tempat tinggal, tingkat glukosa rata-rata, BMI, dan status merokok. Label target adalah variabel biner bernama stroke yang menunjukkan apakah pasien pernah mengalami stroke (1) atau tidak (0).

2.2. Praproses Data

Tahapan praproses dilakukan untuk memastikan data bersih dan siap digunakan oleh algoritma pembelajaran mesin. Langkah-langkahnya meliputi (a) menghapus kolom tidak relevan: kolom id dihapus karena tidak memberikan nilai prediktif, (b) imputasi data hilang: kolom bmi memiliki 201 nilai kosong, yang diisi menggunakan median karena robust terhadap outlier, (c) e*ncoding data categorikal*: *label encoding* untuk kolom biner: *gender, ever_married, residence_type*; *one-hot encoding* untuk kolom *work_type dan smoking_status*, (d) menghapus outlier: nilai kategori *other* pada kolom gender dihapus karena jumlahnya sangat kecil, (e) menangani ketidakseimbangan kelas: kelas target tidak seimbang (stroke = 249, non-stroke = 4861) maka digunakan metode SMOTE (*Synthetic Minority Oversampling Technique*) untuk menyeimbangkan jumlah sampel antar kelas, (f) normalisasi: fitur numerik seperti *age, avg_glucose_level,* dan *bmi* dinormalisasi menggunakan *StandardScaler* agar memiliki distribusi dengan rata-rata 0 dan standar deviasi 1.

2.3. Pengolahan Data

SMOTEENN (*Synthetic Minority Oversampling Technique and Edited Nearest Neighbor*) adalah teknik pengolahan data yang digunakan untuk mengatasi ketidakseimbangan data (*imbalanced data*), yang sering terjadi ketika jumlah sampel

dalam kelas minoritas jauh lebih sedikit dibandingkan kelas mayoritas. SMOTEENN merupakan gabungan dari dua metode yang saling melengkapi, SMOTE (*Synthetic Minority Oversampling Technique*) adalah metode *oversampling* yang menghasilkan data sintetis untuk kelas minoritas. Teknik ini bekerja dengan membuat sampel baru di sepanjang garis antara data minoritas yang ada, alih-alih hanya menduplikasi data.

Hal ini membantu meningkatkan jumlah data di kelas minoritas secara realistis tanpa mengubah distribusi data secara signifikan. dan ENN (*Edited Nearest Neighbor*) adalah teknik *undersampling* yang digunakan setelah proses SMOTE.

ENN berfungsi untuk membersihkan data dengan menghapus sampel dari kelas mayoritas yang salah diklasifikasikan oleh algoritma *Decision Tree* dan XGBoost berdasarkan tetangganya (Herlistiono and Violina 2024). Ini membantu mengurangi noise dan memastikan bahwa data yang digunakan untuk pelatihan lebih berkualitas. Setelah proses ini, dataset menjadi lebih seimbang, meningkatkan kemampuan model dalam mendeteksi kasus pada kelas minoritas.

$$\boldsymbol{x_{new} = x_i + \lambda(x_k - x_i)}$$

Keterangan:
x_i : Sampel minoritas asli
x_k : Tetangga terdekat xix_ixi
λ : Bilangan acak antara 0 dan 1

2.4. Pembagian Data

Setelah praproses dan seleksi fitur, data dibagi menjadi dua bagian yaitu *Training Set*: 80% dari total data, digunakan untuk melatih model dan *Testing Set*: 20% dari total data, digunakan untuk menguji akurasi model. Pembagian dilakukan secara stratified agar distribusi kelas tetap proporsional.

2.5. Pembangunan Model

Empat algoritma pembelajaran mesin digunakan dalam penelitian ini: *AdaBoost (Adaptive Boosting)*, *Gradient Boosting Machines (GBM)*, *LightGBM (Light Gradient Boosting Machine)*, *XGBoost (eXtreme Gradient Boosting).*

Setiap algoritma dilatih menggunakan training set dan dibandingkan kinerjanya berdasarkan metrik evaluasi pada testing set.

2.6. Evaluasi Model

Evaluasi model dilakukan menggunakan *Confusion Matrix* dan metrik berikut: akurasi (*Accuracy*) sebagai rasio prediksi benar terhadap total prediksi. Presisi dan *Recall* sebagai mengukur performa dalam mengidentifikasi kasus stroke. F1-Score untuk harmonik rata-rata dari presisi dan recall. ROC Curve dan AUC digunakan untuk mengukur *trade-off* antara TPR dan FPR.

3. HASIL DAN PEMBAHASAN

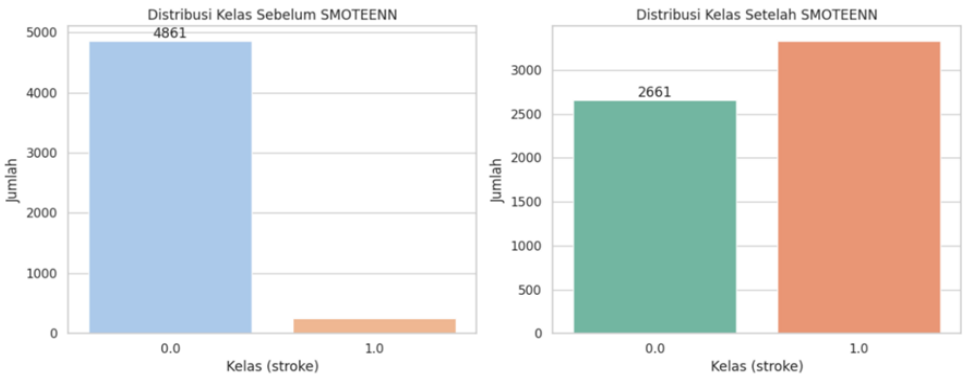
3.1 Hasil Penelitian

Dari dataset prediksi stroke yang disediakan, terdapat 5110 responden dengan berbagai kondisi. Dari 5110 responden tersebut, terdapat 12 atribut yang mencakup: jenis kelamin, usia, hipertensi, penyakit jantung, status perkawinan, jenis pekerjaan, tipe hunian, kadar glukosa rata-rata, BMI, status merokok, dan stroke. Dari 12 atribut yang ada, atribut “id” diabaikan karena tidak berpengaruh pada hasil dapat dilihat pada tabel 1.

Tabel 1. Deskripsi *Dataset*

Atribut	***Description***
ID	Nomor id pasien
Gender	Jenis kelamin pasien
Age	Usia pasien
Hypertension	0 – tidak hipertensi 1 – hipertensi
Heart disease	0 – tidak memiliki riwayat penyakit jantung 1 – memiliki riwayat penyakit jantung
Marital status	Menikah atau tidak menikah
Work-type	Jenis pekerjaan pasien
Residence area	Wilayah tempat tinggal pasien
Avg-glukose	Rata-rata level glukosa dalam darah yang diukur setelah makan
BMI	*Body Mass Index* pasien
Smoking status	Status merokok pasien
Stroke status	0 – tidak stroke 1 – stroke

a. Klasifikasi Hasil

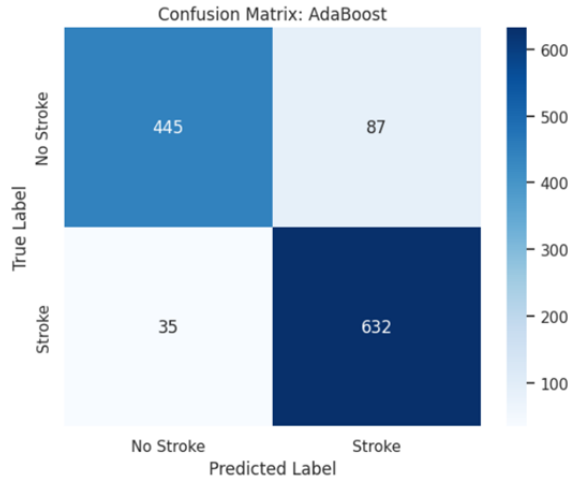


Gambar 2. Proporsi Data Stroke Sebelum dan Sesudah SMOTEENN

Pada gambar 2 menampilkan distribusi kelas target stroke sebelum dan sesudah diterapkannya metode penyeimbangan data SMOTEENN (*Synthetic Minority Over-sampling Technique and Edited Nearest Neighbors*). Pada grafik sebelah kiri yang menunjukkan kondisi sebelum dilakukan penyeimbangan, terlihat adanya ketimpangan distribusi kelas yang sangat signifikan. Kelas 0 (tidak mengalami stroke) memiliki jumlah data sebanyak 4.861, sedangkan kelas 1 (mengalami stroke) hanya berjumlah 249. Ketidakseimbangan ini menyebabkan model *machine learning* yang dilatih dengan data tersebut cenderung bias terhadap kelas mayoritas dan mengabaikan kelas minoritas, padahal dalam konteks prediksi medis seperti stroke, mendeteksi kasus minoritas justru jauh lebih krusial.

Setelah dilakukan penyeimbangan dengan teknik SMOTEENN seperti yang terlihat pada grafik sebelah kanan, distribusi kelas menjadi jauh lebih seimbang. Jumlah data pada kelas 0 berkurang menjadi 2.661, sementara jumlah data pada kelas 1 meningkat menjadi 3.331. Teknik SMOTE bertugas melakukan oversampling terhadap kelas minoritas dengan menciptakan data sintetis berdasarkan tetangga terdekatnya, sedangkan ENN berperan melakukan undersampling terhadap kelas mayoritas dengan menghapus data yang secara klasifikasi dianggap ambigu atau noise. Kombinasi kedua metode ini tidak hanya menyeimbangkan jumlah data antar kelas, tetapi juga membersihkan data dari potensi kesalahan klasifikasi.

Penyeimbangan data seperti ini sangat penting untuk meningkatkan kinerja model dalam mengenali pola dari kedua kelas secara adil. Dalam konteks klasifikasi medis, hasil ini sangat menguntungkan karena membantu model menjadi lebih sensitif terhadap kasus-kasus stroke yang sebelumnya langka dalam data. Hal ini diharapkan dapat meningkatkan metrik performa model seperti *recall, precision, dan F1-score* pada kelas minoritas, serta mengurangi potensi kesalahan dalam diagnosis yang bisa berdampak serius bagi pasien. Dengan demikian, penerapan SMOTEENN terbukti efektif dalam memperbaiki distribusi data dan mendukung pengembangan model prediktif yang lebih akurat dan andal.



Gambar 3. *Confusion Matrix* AdaBoost

Model klasifikasi yang menggunakan algoritma AdaBoost untuk memprediksi kejadian stroke berdasarkan data medis. Matriks ini memberikan gambaran mengenai performa model dalam mengklasifikasikan dua kategori, yaitu *"No Stroke"* dan *"Stroke"*, baik dari sisi label sebenarnya ***(****True Label****)*** maupun label hasil prediksi model ***(****Predicted Label****)***. Dari hasil gambar 3, terlihat bahwa model memiliki performa yang cukup baik dalam mendeteksi stroke, karena nilai *True Positive* (632) jauh lebih tinggi dibandingkan *False Negative* (35). Begitu pula dalam mendeteksi pasien yang tidak mengalami stroke, model menghasilkan *True Negative* sebesar 445, yang juga cukup signifikan dibandingkan dengan *False Positive* sebesar 87. Skema warna biru dalam heatmap menunjukkan intensitas jumlah prediksi, di mana semakin gelap warnanya, semakin banyak jumlah data pada sel tersebut. Warna tergelap terdapat pada sel *True Positive* (632), menandakan prediksi stroke yang paling banyak dan paling akurat dilakukan oleh model. Secara keseluruhan, *confusion matrix* ini menunjukkan bahwa model AdaBoost bekerja cukup efektif dalam klasifikasi biner untuk kasus deteksi stroke.

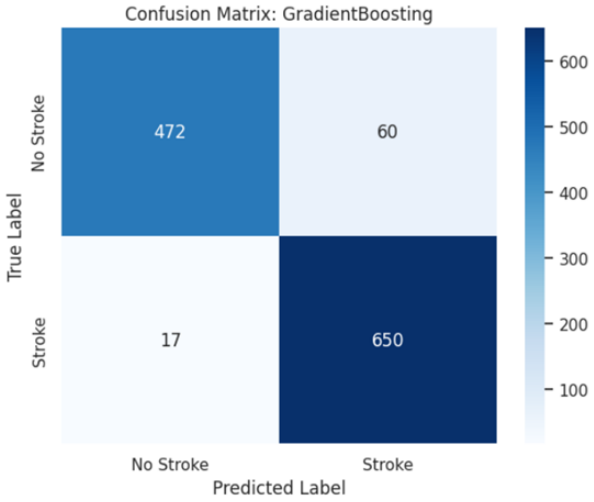
Classification Report:

	precision	recall	f1-score	support
0.0	0.93	0.84	0.88	532
1.0	0.88	0.95	0.91	667
accuracy			0.90	1199
macro avg	0.90	0.89	0.90	1199
weighted avg	0.90	0.90	0.90	1199

Gambar 4. *Classification Report* AdaBoost

Gambar di atas menunjukkan hasil evaluasi kinerja model klasifikasi AdaBoost dalam bentuk *classification report*. Berdasarkan *classification report*, untuk kelas 0.0 *(No Stroke)*, model memiliki *precision* sebesar 93%, *recall* sebesar 84%, dan f1-score sebesar

88% dengan jumlah data *(support)* sebanyak 532. Sementara itu, untuk kelas 1.0 (Stroke), *precision* mencapai angka sempurna yaitu 100%, *recall* sebesar 95%, dan *f1-score* sebesar 91% dari 667 data. Secara keseluruhan, akurasi model berada pada angka 0.898 atau sekitar 89.8%, menunjukkan bahwa hampir 90% dari total prediksi yang dihasilkan oleh model adalah tepat. Nilai rata-rata untuk *precision*, *recall*, dan *f1-score*, baik secara makro maupun tertimbang (weighted), semuanya berada di angka 90%, yang mencerminkan konsistensi dan keseimbangan performa model terhadap kedua kelas. Berdasarkan hasil ini, dapat disimpulkan bahwa algoritma AdaBoost menunjukkan kinerja yang sangat baik dan andal dalam mendeteksi kasus stroke dengan tingkat kesalahan yang relatif rendah.



Gambar 5. *Confusion Matrix Gradient Boosting*

Model klasifikasi *Gradient Boosting* pada gambar 5 menampilkan hasil prediksi model dibandingkan dengan label sebenarnya, yang dibagi dalam dua kategori yaitu *"No Stroke"* dan *"Stroke"*, baik dari sisi label asli *(True Label)* maupun hasil prediksi *(Predicted Label)*. terlihat bahwa model *Gradient Boosting* memiliki performa yang sangat baik, dengan angka *True Positive* dan *True Negative* yang tinggi, yaitu masing-masing 650 dan 472, serta angka *False Negative* yang sangat rendah (17). Ini menunjukkan kemampuan model dalam mengenali kasus stroke secara akurat, yang sangat penting dalam konteks deteksi dini dan penanganan medis. Warna dalam heatmap ini menunjukkan intensitas jumlah kasus, dengan warna biru tua pada sel *True Positive* menandakan bahwa sebagian besar data berhasil diprediksi dengan benar sebagai *"Stroke"*.

Secara keseluruhan, model *Gradient Boosting* tampak memiliki akurasi tinggi dan kesalahan minimal, terutama dalam mendeteksi pasien yang benar-benar mengalami stroke, yang membuatnya sangat efektif dan dapat diandalkan dalam skenario klasifikasi medis seperti ini.

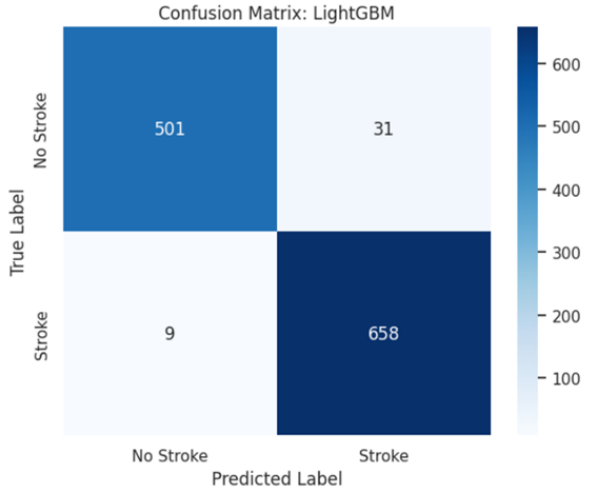
Classification Report:

	precision	recall	f1-score	support
0.0	0.97	0.89	0.92	532
1.0	0.92	0.97	0.94	667
accuracy			0.94	1199
macro avg	0.94	0.93	0.93	1199
weighted avg	0.94	0.94	0.94	1199

Gambar 6. *Classification Report Gradient Boosting*

Gambar di atas merupakan *classification report* yang menampilkan hasil evaluasi kinerja model *Gradient Boosting* dalam mendeteksi dua kelas, yaitu 0.0 (No Stroke) dan 1.0 (Stroke), berdasarkan metrik *precision, recall, f1-score*, dan *support*. Untuk kelas 0.0, model mencapai *precision* sebesar 0.97, yang berarti 97% dari prediksi *"No Stroke"* adalah benar. *Recall* untuk kelas ini sebesar 0.89, menunjukkan bahwa 89% dari kasus nyata *"No Stroke"* berhasil dikenali oleh model. Kombinasi *precision* dan *recall* menghasilkan nilai *f1-score* sebesar 0.92, dengan total 532 data aktual untuk kelas ini. Sementara itu, untuk kelas 1.0 (Stroke), *precision model* sebesar 0.92, *recall* sangat tinggi yaitu 0.97, dan *f1-score* mencapai 0.94 dari 667 data aktual, menunjukkan bahwa model sangat efektif dalam mendeteksi kasus stroke.

Secara keseluruhan, akurasi model mencapai 0.94 atau 94%, yang berarti dari 1199 data yang diuji, sekitar 94% prediksi model sesuai dengan label sebenarnya. Nilai rata-rata makro untuk *precision, recal*l, dan *f1-score* masing-masing adalah 0.94, 0.93, dan 0.93. Sementara itu, rata-rata tertimbang *(weighted average)* yang mempertimbangkan jumlah data per kelas menghasilkan nilai yang sama, yaitu *precision* 0.94, *recall* 0.94, dan *f1-score* 0.94. Hasil ini menunjukkan bahwa model tidak hanya akurat secara keseluruhan, tetapi juga seimbang dalam kinerjanya terhadap kedua kelas, menjadikannya andal dan efektif, terutama dalam konteks deteksi medis seperti identifikasi kasus stroke.



Gambar 7. *Confusion Matrix LightGBM*

Pada gambar 7 menampilkan *confusion matrix* dari model *LightGBM* yang digunakan untuk klasifikasi dua kelas, yaitu *No Stroke* dan Stroke. Dari matriks tersebut terlihat bahwa model berhasil mengklasifikasikan 501 data *No Stroke* dengan benar *(True Negative)* dan hanya melakukan kesalahan pada 31 data *No Stroke* yang diprediksi sebagai Stroke *(False Positive)*. Sementara itu, untuk data dengan label sebenarnya Stroke, model menunjukkan kinerja yang sangat baik dengan mengklasifikasikan 658 data Stroke secara akurat *(True Positive)* dan hanya melakukan kesalahan pada 9 data Stroke yang diprediksi sebagai *No Stroke (False Negative)*.

Dengan jumlah kesalahan prediksi yang sangat kecil baik pada kelas *No Stroke* maupun Stroke, dapat disimpulkan bahwa model *LightGBM* memiliki kinerja klasifikasi yang sangat tinggi dan presisi yang baik, khususnya dalam mendeteksi kasus Stroke yang umumnya menjadi fokus utama dalam aplikasi medis. Akurasi tinggi pada kedua kelas juga menunjukkan bahwa model ini tidak mengalami bias signifikan terhadap salah satu kelas, menjadikannya andal dan layak diterapkan dalam sistem deteksi stroke berbasis data.

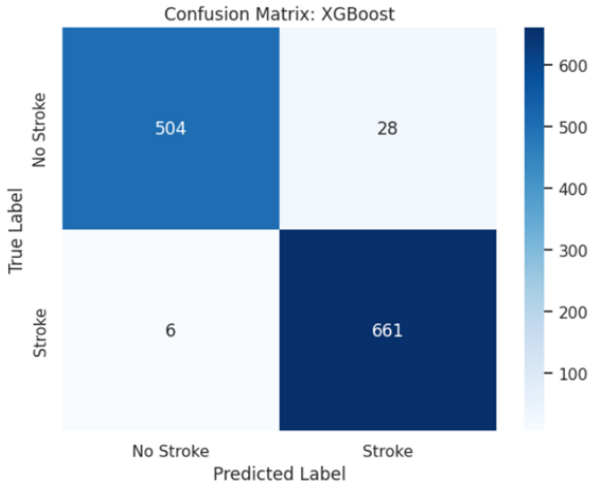
Classification Report:

	precision	recall	f1-score	support
0.0	0.98	0.94	0.96	532
1.0	0.96	0.99	0.97	667
accuracy			0.97	1199
macro avg	0.97	0.96	0.97	1199
weighted avg	0.97	0.97	0.97	1199

Gambar 8. *Classification Report LightGBM*

Gambar di atas menunjukkan classification report dari model klasifikasi *LightGBM* yang mengevaluasi performa dalam mendeteksi dua kelas, yaitu 0.0 *(No Stroke)* dan 1.0 (Stroke). Untuk kelas 0.0, model menunjukkan *precision* sebesar 0.98, artinya 98% prediksi *"No Stroke"* adalah benar. *Recall*-nya mencapai 0.94, yang berarti 94% dari seluruh kasus nyata *"No Stroke"* berhasil dikenali oleh model. *F1-score* untuk kelas ini tercatat sebesar 0.96, dengan jumlah data aktual sebanyak 532. Untuk kelas 1.0 (Stroke), *precision model* sebesar 0.96, *recall* sangat tinggi yaitu 0.99, dan *f1-score* sebesar 0.97, dari 667 data aktual, menunjukkan kemampuan yang sangat baik dalam mendeteksi kasus stroke.

Secara keseluruhan, akurasi model mencapai 0.97 atau 97%, yang berarti 97% dari total 1199 data diklasifikasikan dengan benar. Rata-rata makro (macro avg) untuk *precision, recall,* dan *f1-score* masing-masing adalah 0.97, 0.96, dan 0.97, sedangkan rata-rata tertimbang *(weighted avg*) menunjukkan angka yang sama tinggi, yaitu *precision* dan *recall* sebesar 0.97, serta *f1-score* sebesar 0.97. Nilai-nilai ini mencerminkan performa yang sangat konsisten dan seimbang antara kedua kelas. Dengan akurasi yang sangat tinggi dan performa metrik lainnya yang unggul, model ini terbukti sangat efektif dan andal dalam melakukan klasifikasi, terutama dalam konteks deteksi medis seperti prediksi risiko stroke.



Gambar 9. *Confusion Matrix XGBoost*

Pada gambar 9 menunjukkan *confusion matrix* dari model *XGBoost* yang digunakan untuk melakukan klasifikasi terhadap dua kelas, yaitu *No Stroke* dan Stroke. Berdasarkan matriks tersebut, model berhasil mengklasifikasikan 504 data *No Stroke* secara benar *(True Negative)* dan hanya melakukan kesalahan pada 28 data yang seharusnya *"No Stroke"* namun diprediksi sebagai "Stroke" *(False Positive)*. Sementara itu, untuk data yang benar-benar merupakan kasus "Stroke", model mengklasifikasikan dengan sangat akurat sebanyak 661 data *(True Positive)*, dan hanya 6 data Stroke yang salah diklasifikasikan sebagai *"No Stroke" (False Negative).*

Dengan jumlah kesalahan prediksi yang sangat rendah di kedua kelas, *confusion matrix* ini menunjukkan bahwa model *XGBoost* memiliki kinerja klasifikasi yang sangat baik dan seimbang, terutama dalam mendeteksi kasus stroke yang sangat krusial. Keseimbangan antara ketepatan dalam memprediksi pasien yang sehat maupun yang berisiko stroke menjadikan model ini sangat andal dan potensial untuk digunakan dalam sistem pendukung keputusan medis berbasis *machine learning*.

Classification Report:

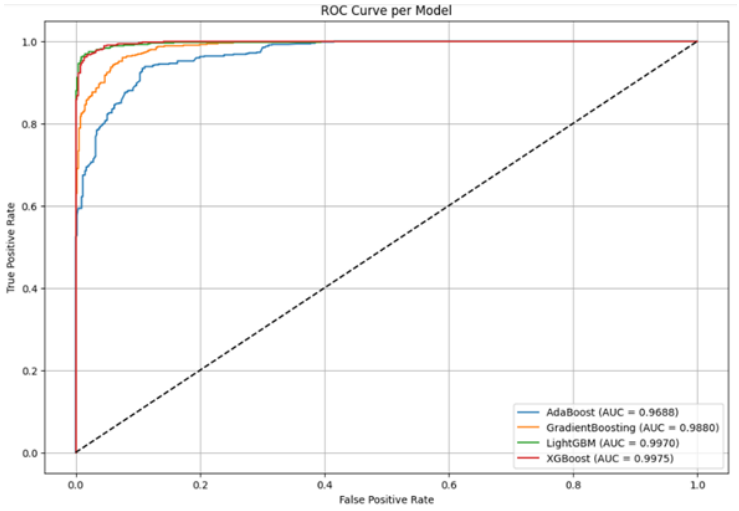
	precision	recall	f1-score	support
0.0	0.99	0.95	0.97	532
1.0	0.96	0.99	0.97	667
accuracy			0.97	1199
macro avg	0.97	0.97	0.97	1199
weighted avg	0.97	0.97	0.97	1199

Gambar 10. *Classification Report XGBoost*

Gambar di atas menunjukkan *classification report* dari model *XGBoost* yang mengevaluasi performanya terhadap dua kelas, yaitu kelas 0.0 *(No Stroke)* dan kelas 1.0 (Stroke). Untuk kelas 0.0, model menunjukkan nilai *precision* sebesar 0.99, artinya dari semua prediksi *"No Stroke"*, 99% benar. Nilai *recall* sebesar 0.95 menunjukkan bahwa 95% dari seluruh data aktual *"No Stroke"* berhasil diidentifikasi dengan tepat. Nilai *f1-score* untuk kelas ini adalah 0.97, yang menunjukkan keseimbangan yang sangat baik antara *precision* dan *recall*, dengan jumlah data aktual sebanyak 532.

Untuk kelas 1.0 (Stroke), *precision model* adalah 0.96, *recall* mencapai 0.99, dan *f1-score* juga 0.97, berdasarkan total 667 data aktual. Nilai *recall* yang tinggi pada kelas stroke sangat penting karena menunjukkan kemampuan model dalam mendeteksi sebagian besar kasus stroke dengan benar. Secara keseluruhan, akurasi model adalah 0.97, atau 97%, dari total 1199 sampel. Baik *macro average* maupun *weighted average* untuk *precision*, *recall*, dan *f1-sc*ore masing-masing tercatat sebesar 0.97, menunjukkan bahwa performa model sangat konsisten dan seimbang dalam mengklasifikasikan kedua kelas.

Secara umum, hasil ini menandakan bahwa model memiliki kinerja klasifikasi yang sangat tinggi, dengan kesalahan yang minimal dan kemampuan yang sangat baik dalam membedakan antara pasien yang berisiko stroke dan yang tidak. Ini menjadikan model ini sangat potensial untuk diterapkan dalam sistem prediksi kesehatan berbasis data.



Gambar 11. *ROC Curve* dan *AUC Model*

Gambar di atas menunjukkan kurva *ROC (Receiver Operating Characteristic)* dari empat model klasifikasi yang digunakan untuk mendeteksi kasus stroke, yaitu *AdaBoost*, *Gradient Boosting*, *LightGBM*, dan *XGBoost*. Kurva ROC menggambarkan hubungan antara *true positive rate* (sensitivitas) dan *false positive rate* (1-spesifisitas) pada berbagai ambang klasifikasi. Semakin dekat kurva ROC ke pojok kiri atas, semakin baik performa model tersebut dalam membedakan antara dua kelas.

Dari gambar, terlihat bahwa semua model memiliki performa yang sangat baik, dengan area di bawah kurva (AUC) yang tinggi. *XGBoost* menunjukkan performa terbaik dengan nilai AUC sebesar 0.9975, yang mendekati sempurna. *LightGBM* menyusul dengan AUC sebesar 0.9970, disusul oleh *Gradient Boosting* dengan AUC 0.9880, dan terakhir *AdaBoost* dengan AUC 0.9688. Nilai AUC yang tinggi menunjukkan bahwa keempat model memiliki kemampuan yang sangat baik dalam membedakan antara pasien yang mengalami stroke dan yang tidak.

Secara keseluruhan, kurva ROC ini menunjukkan bahwa model *XGBoost* sangat unggul dalam hal diskriminasi klasifikasi, dengan performa yang jauh melebihi garis diagonal *(baseline)* yang merepresentasikan prediksi acak. Hasil ini mendukung kesimpulan bahwa algoritma berbasis *boosting*, khususnya *XGBoost* sangat cocok digunakan dalam sistem prediksi medis yang membutuhkan sensitivitas tinggi.

b. Perbandingan Model

Tabel Perbandingan Model (dalam %):

	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
0	AdaBoost	89.82	87.90	94.75	91.20	96.88
1	GradientBoosting	93.58	91.55	97.45	94.41	98.80
2	LightGBM	96.66	95.50	98.65	97.05	99.70
3	XGBoost	97.16	95.94	99.10	97.49	99.75

Gambar 12. Tabel Perbandingan Model

Pada gambar 12 menampilkan tabel perbandingan kinerja empat model klasifikasi yang digunakan untuk mendeteksi stroke, yaitu *AdaBoost, Gradient Boosting, LightGBM, dan XGBoost*, berdasarkan beberapa metrik evaluasi utama: akurasi, *precision, recall*, *F1-score*, dan *AUC (Area Under the Curve)*.

Model *AdaBoost* menunjukkan performa terendah di antara keempat model, dengan akurasi sebesar 89.82%, *precision* 87.90%, *recall* 94.75%, *F1-score* 91.20%, dan AUC 96.88%. Meskipun *recall-nya* cukup tinggi, *precision* yang relatif lebih rendah menunjukkan adanya prediksi positif palsu yang lebih banyak dibandingkan model lain.

Model *Gradient Boosting* memperlihatkan peningkatan signifikan dengan akurasi 93.58%, *precision* 91.55%, *recall* 97.45%, *F1-score* 94.41%, dan AUC 98.80%. Ini menunjukkan bahwa model ini lebih seimbang dalam menangani prediksi positif dan negatif, serta memiliki kemampuan deteksi yang baik.

LightGBM bahkan menunjukkan hasil yang lebih tinggi dengan akurasi 96.66%, *precision* 95.50%, *recall* 98.65%, *F1-score* 97.05%, dan AUC 99.70%. Model ini sangat efisien dalam mendeteksi stroke, dengan keseimbangan luar biasa antara precision dan recall.

Model terbaik berdasarkan keseluruhan metrik adalah *XGBoost*, dengan akurasi 97.16%, *precision* 95.94%, *recall* 99.10%, *F1-score* 97.49%, dan AUC 99.75%. Nilai-nilai ini menandakan bahwa *XGBoost* memiliki kemampuan prediksi yang sangat kuat, minim kesalahan klasifikasi, dan sangat cocok diterapkan dalam konteks medis yang menuntut akurasi tinggi.

3.2 Pembahasan

Berdasarkan hasil keseluruhan dari evaluasi model klasifikasi untuk deteksi stroke, terlihat bahwa algoritma *boosting modern* yaitu *XGBoost* menunjukkan performa yang paling unggul dibandingkan dengan AdaBoost, Gradient Boosting dan *LightGBM*. Hal ini ditunjukkan melalui nilai akurasi, *precision*, *recall*, *F1-score*, dan AUC yang sangat tinggi, pada *XGBoost* yang mencapai akurasi 97.16% dan AUC 99.75%. Sementara itu, meskipun AdaBoost masih mampu memberikan hasil yang cukup baik, performanya berada di bawah model-model lainnya, terutama dalam hal precision. Secara umum, *XGBoost* terbukti sangat efektif dalam mendeteksi kasus stroke dengan tingkat kesalahan yang sangat rendah, menjadikannya pilihan yang lebih andal untuk aplikasi klasifikasi medis berbasis data.

Hal ini juga diperkuat dengan penelitian sebelumnya yang terdapat pada tabel 2 dibawah ini:

Tabel 2. Perbandingan Model Penelitian Sebelumnya

Penelitian Berkaitan	Tahun	Model	Akurasi
Our Model	2025	XGBoost	97.16%
(Augi and Sultan 2024)	2024	XGBoost	96.45%
(Wulandari, Isro'Mukti, and Susanti 2024)	2024	KNN	88.46%
(Hasibuan et al. 2022)	2022	Naïve Bayes-PSO	95.02%
(Fadillah Hermawan, Rakhmat Umbara, and Kasyidi 2022)	2022	CART	89.83%
(Azhar, Firdausy, and Amelia 2022)	2022	DNN	96%
(Maskuri, Sukerti, and Herdian Bhakti 2022)	2022	KNN	95%
(Bugis 2022)	2022	Naïve Bayes	93.94%
(Iskandar, Ernawati, and Widiastiwi 2022)	2022	Random Forest	95.02%
(Rahim, Sunyoto, and Arief 2022)	2022	XGBoost	96%

4. KESIMPULAN

Penelitian ini berhasil mengembangkan model prediktif untuk deteksi dini stroke dengan memanfaatkan berbagai algoritma *machine learning*, yaitu *AdaBoost*, *Gradient Boosting*, *LightGBM*, dan *XGBoost*. Dari keempat model yang diuji, *XGBoost* menunjukkan performa terbaik dengan akurasi sebesar 97,16% dan AUC sebesar 99,75%, menjadikannya sangat andal dalam mengidentifikasi kasus stroke dengan tingkat kesalahan minimal. Keberhasilan model tidak hanya ditentukan oleh pemilihan algoritma, tetapi juga oleh tahapan praproses data yang cermat seperti penanganan ketidakseimbangan data menggunakan SMOTEENN, normalisasi, serta seleksi fitur yang relevan. Penelitian ini menunjukkan bahwa pendekatan komprehensif yang mencakup keseluruhan siklus pemodelan dapat menghasilkan sistem prediksi yang sangat efektif dalam konteks medis. Selain itu, hasil ini memperkuat potensi machine learning sebagai alat bantu dalam pengambilan keputusan klinis yang proaktif dan berbasis data. Meski demikian, validasi lanjutan pada data klinis nyata dan integrasi dengan sistem informasi rumah sakit menjadi langkah penting untuk implementasi model secara luas. Penelitian ini diharapkan menjadi pijakan awal bagi pengembangan sistem deteksi stroke yang lebih akurat, transparan, dan dapat diadopsi secara praktis dalam layanan kesehatan digital.

Untuk pengembangan di masa depan, peneliti merekomendasikan perluasan penelitian ke arah integrasi dengan data elektronik rekam medis *(Electronic Health Records/EHR)*, pengembangan sistem rekomendasi interaktif untuk intervensi medis, dan pemanfaatan teknik explainable AI (XAI) untuk meningkatkan transparansi model. Selain itu, kolaborasi antara peneliti teknologi dan tenaga medis sangat penting untuk memastikan bahwa model yang dikembangkan benar-benar dapat diadopsi dalam praktik klinis secara aman dan efektif.

REFERENCES

Augi, Esraa H., and Almabruk Sultan. 2024. "The Early Warning Signs of a Stroke: An Approach Using Machine Learning Predictions." *Journal of Computer and Communications* 12(06): 59–71. doi:10.4236/jcc.2024.126005.

Azhar, Yufis, Aidia Khoiriyah Firdausy, and Putri Juli Amelia. 2022. "Perbandingan Algoritma Klasifikasi Data Mining Untuk Prediksi Penyakit Stroke." *SINTECH (Science and Information Technology) Journal* 5(2): 191–97. doi:10.31598/sintechjournal.v5i2.1222.

Bugis, Haris. 2022. "Metode Naïve Bayes Untuk Memprediksi Penyakit Stroke." *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan)* 6(1): 8–14. doi:10.47970/siskom-kb.v6i1.317.

Chung, Chen Chih, Emily Chia Yu Su, Jia Hung Chen, Yi Tui Chen, and Chao Yang Kuo. 2023. "XGBoost-Based Simple Three-Item Model Accurately Predicts Outcomes of Acute Ischemic Stroke." *Diagnostics* 13(5): 1–13. doi:10.3390/diagnostics13050842.

Fadillah Hermawan, Agiel, Fajri Rakhmat Umbara, and Fatan Kasyidi. 2022. "MIND (Multimedia Artificial Intelligent Networking Database Prediksi Awal Penyakit Stroke Berdasarkan Rekam Medis Menggunakan Metode Algoritma CART(Classification and Regression Tree)." *Journal MIND Journal | ISSN* 7(2): 151–64. https://doi.org/10.26760/mindjournal.v7i2.151-164.

Gupta, Aakanshi, Nidhi Mishra, Nishtha Jatana, Shaily Malik, Khaled A. Gepreel, Farwa Asmat, and Sachi Nandan Mohanty. 2024. "Predicting Stroke Risk: An Effective Stroke Prediction Model Based on Neural Networks." *Journal of Neurorestoratology* 13(1): 100156. doi:10.1016/j.jnrt.2024.100156.

Hasibuan, M Said, Devi Fransisca, Jurusan Magister, Teknik Informatika, and Fakultas Ilmu Komputer. 2022. "Penggunaan Algoritma Naive Bayes Dan Particle Swarm Optimization (PSO) Untuk Mendeteksi Stroke." : 109–18.

Herlistiono, Iwa Ovyawan, and Sriyani Violina. 2024. "Model Prediksi Risiko Stroke Menggunakan Machine Learning." *INTECOMS: Journal of Information Technology and Computer Science* 7(4): 1230–38. doi:10.31539/intecoms.v7i4.10942.

Indah Werdiningsih, Endah Purwanti, Iin Mardiyana, Arum Tiyas Handayani, Kharristantie Sekarlangit Suryadewi, Endang Nurjanah, Fildzah Akhlaqulkarimah, Naurah Hedy Pramiyas, and Fakhrana Almas Syah Yahrani. 2023. "Analisis Prediksi Stroke Menggunakan Pendekatan Decision Tree Dengan Seleksi Fitur Dan Neural Network." *Jurnal Sistem Cerdas* 6(3): 213–21. doi:10.37396/jsc.v6i3.310.

Iskandar, Nur Aliffiyanti, Iin Ernawati, and Yuni Widiastiwi. 2022. "Klasifikasi Diagnosis Penyakit Stroke Dengan Menggunakan Metode Random Forest." *Seminak Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*: 432–41. https://conference.upnvj.ac.id/index.php/senamika/article/view/2190.

Maskuri, Muhammad Naja, Kadek Sukerti, and R M Herdian Bhakti. 2022. "Penerapan Algoritma K-Nearest Neighbor (KNN) Untuk Memprediksi Penyakit Stroke Stroke Desease Predict Using KNN Algorithm." *Jurnal Ilmiah Intech : Information Technology Journal of UMUS* 4(1): 130–40.

Putra, Leonardus Sandy Ade, Eka Kusumawardhani, Putranty Widha Nugraheni, Lalak Tarbiyatun Nasyin Maleiva, and Vincentius Abdi Gunawan. 2022. "Sistem Identifikasi Dini Penyakit Stroke Dengan Menggunakan Jaringan Syaraf Tiruan Perambatan Balik." *Jurnal Teknologi Informasi: Jurnal Keilmuan dan Aplikasi Bidang Teknik Informatika* 16(2): 145–57. doi:10.47111/jti.v16i2.5096.

Rahim, Abd Mizwar A, Andi Sunyoto, and Muhammad Rudyanto Arief. 2022. "Stroke Prediction Using Machine Learning Method with Extreme Gradient Boosting Algorithm." *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer* 21(3): 595–606. doi:10.30812/matrik.v21i3.1666.

Salsabila, Mutiara, Novi Susanti, Niken Natani Sabilla, Henny Irene, Natalia Hulu, Program Studi, Ilmu Kesehatan, et al. 2024. "Analisis Kebijakan Kesehatan Tentang Pencegahan Dan Penanganan Stroke Di Indonesia." 3(2): 931–38.

Swathi Priyadarshini, T., and Mohd Abdul Hameed. 2024. "Developing Heart Stroke Prediction Model Using Deep Learning with Combination of Fixed Row Initial Centroid Method with Navie Bayes, Decision Tree, and Artificial Neural Network." *Measurement: Sensors* 34(May): 101237. doi:10.1016/j.measen.2024.101237.

Wulandari, Serin, Yogi Isro'Mukti, and Tri Susanti. 2024. "Optimalisasi Prediksi Penyakit Stroke Menggunakan Algoritma Deep Learning." *JATI (Jurnal Mahasiswa Teknik Informatika)* 8(2): 1826–33. doi:10.36040/jati.v8i2.9256.

Zuama, Robi Aziz, Syaifur Rahmatullah, and Yuri Yuliani. 2022. "Analisis Performa Algoritma Machine Learning Pada Prediksi Penyakit Cerebrovascular Accidents." *Jurnal Media Informatika Budidarma* 6(1): 531. doi:10.30865/mib.v6i1.3488.